The mass and anisotropy profiles of galaxy clusters from the projected phase space density: testing the method on simulated data

Radosław Wojtak,[1] Ewa L. Łokas,[1] Gary A. Mamon[2,3] and Stefan Gottlöber[4]

[1] *Nicolaus Copernicus Astronomical Center, Bartycka 18, 00-716 Warsaw, Poland*
[2] *Institut d'Astrophysique de Paris (UMR 7095: CNRS and Université Pierre & Marie Curie), 98 bis Bd Arago, F-75014 Paris, France*
[3] *Astrophysics 7 BIPAC, University of Oxford, Keble Rd, OX1 3RH, Oxford, UK*
[4] *Astrophysikalisches Institut Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany*

27 June 2009

ABSTRACT

We present a new method of constraining the mass and velocity anisotropy profiles of galaxy clusters from kinematic data. The method is based on a model of the phase space density which allows the anisotropy to vary with radius between two asymptotic values. The characteristic scale of transition between these asymptotes is fixed and tuned to a typical anisotropy profile resulting from cosmological simulations. The model is parametrized by two values of anisotropy, at the centre of the cluster and at infinity, and two parameters of the NFW density profile, the scale radius and the scale mass. In order to test the performance of the method in reconstructing the true cluster parameters we analyze mock kinematic data for 20 relaxed galaxy clusters generated from a cosmological simulation of the standard ΛCDM model. We use Bayesian methods of inference and the analysis is carried out following the Markov Chain Monte Carlo approach. The parameters of the mass profile are reproduced quite well, but we note that the mass is typically underestimated by 15 percent, probably due to the presence of small velocity substructures. The constraints on the anisotropy profile for a single cluster are in general barely conclusive. Although the central asymptotic value is determined accurately, the outer one is subject to significant systematic errors caused by substructures at large clustercentric distance. The anisotropy profile is much better constrained if one performs joint analysis of at least a few clusters. In this case it is possible to reproduce the radial variation of the anisotropy over two decades in radius inside the virial sphere.

Key words: galaxies: clusters: general – galaxies: kinematics and dynamics – cosmology: dark matter

1 INTRODUCTION

Kinematic data play an important role in dynamical studies of galaxy clusters. They offer a unique possibility to constrain simultaneously the total (dark and luminous) mass profile and the orbital structure of galaxies which is commonly quantified by the anisotropy of velocity dispersion tensor (Binney & Tremaine 2008). The main challenge of this approach is the fact that both factors, the mass and the anisotropy profile, are interconnected. In the particular case of commonly used Jeans analysis of velocity dispersion profile this leads to a well known problem of the mass-anisotropy degeneracy that hampers the constraining power of the data (e.g. Binney & Mamon 1982; Merritt 1987; Mer-

rifield & Kent 1990). It turns out that in general there are two ways to break this degeneracy: one can either combine the results with other constraints on the mass profile or use a dynamical model going beyond the Jeans equation.

The first solution has been widely applied in the literature. Promising constraints on the anisotropy of galactic orbits in clusters have been obtained by combining the results from velocity dispersion profiles with the mass constraints from X-ray gas (e.g. Benatov et al. 2006; Hwang & Lee 2008) or lensing data (e.g. Natarajan & Kneib 1996). A different approach was adopted by Biviano & Katgert (2004) who argued for an isotropic velocity dispersion tensor of early type galaxies in clusters from the ESO Nearby Abell Cluster Survey (ENACS). Assuming this property they inferred

the total mass profile from velocity dispersions of ellipticals and used this result to constrain the anisotropy profile of late type galaxies. This approach was an implementation of the anisotropy inversion algorithm (Binney & Mamon 1982; Solanes & Salvador-Solé 1990).

The second method of breaking the mass-anisotropy degeneracy relies on the extension of the classical Jeans formalism. The first natural step in this field is to consider the fourth velocity moment or kurtosis (Merrifield & Kent 1990). The method based on the joint fitting of velocity moments (dispersion and kurtosis) was shown to provide constraints both on the mean anisotropy of a system and the parameters of the mass profile (Łokas 2002; Łokas & Mamon 2003; Sanchis, Łokas & Mamon 2004; Łokas et al. 2006). These results confirmed the idea that any attempt to infer the anisotropy profile from kinematic data of spherical systems must be preceded by the construction of a detailed dynamical model. In general there are two ways to achieve a desired complexity of a model. One is the so-called Schwarzschild modelling (Schwarzschild 1979) in which one considers a superposition of base orbits defined in the integral space (e.g. Merritt & Saha 1993; Gerhard et al. 1998; Chanamé, Kleyna & van der Marel 2008). Another one, which we adopt in this work, is to provide a properly parametrized form for the phase space density.

There were several studies devoted to the analysis of kinematic data in terms of the phase space density. An important step in this field was made by Kent & Gunn (1982) who used a family of simple analytical models of the distribution function to analyze the data for the Coma cluster. Van der Marel et al. (2000) obtained constraints on the anisotropy of 16 galaxy clusters from the CNOC1 (Canadian Network for Observational Cosmology) cluster redshift survey. A conceptually similar analysis was carried out by Mahdavi & Geller (2004) for galaxy groups and clusters. In all cases a constant anisotropy was assumed which does not reproduce well the results of cosmological simulations where the dependence of the anisotropy on radius is usually seen (e.g. Mamon & Łokas 2005; Wojtak et al. 2005; Ascasibar & Gottlöber 2008). Since the anisotropy profile has recently become a subject of growing interest, it appears reasonable to generalize the above methods so that both the mass and the anisotropy profiles may be inferred from the data. This implies that one has to deal with an anisotropic model of the phase space density which accounts for its radial variation. Quite recently several models satisfying this requirement have been proposed. The anisotropy profile is specified by a proper parametrization of the angular momentum part of the distribution function (Wojtak et al. 2008) or the augmented density (Van Hese, Baes & Dejonghe 2009). The purpose of the present work was to adopt the approach of Wojtak et al. (2008) to the Bayesian analysis of kinematic data and to test on mock data sets how well the mass and anisotropy profiles are reproduced.

The paper is organized as follows. In the first section we introduce a model of the phase space density and discuss its projection on to the plane of sky. In section 2 we describe mock kinematic data of galaxy clusters generated from a cosmological simulation. Section 3 provides technical details on the Monte Carlo Markov Chain analysis and section 4 presents the results. The discussion follows in section 5.

2 THE PHASE SPACE DENSITY

Any spherical system in equilibrium embedded in a fixed gravitational potential is described completely by the distribution function which depends on the phase space coordinates through the binding energy E and the absolute value of the angular momentum L. In this work, we use the model of the distribution function recently proposed by Wojtak et al. (2008) which was shown to recover spherically averaged phase space distribution of dark matter particles in simulated cluster-size haloes. The main idea of this approach lies in the assumption that the distribution function is separable in energy and angular momentum, i.e. $f(E, L) = f_E(E) f_L(L)$. The angular momentum part $f_L(L)$ is given by an analytical ansatz motivated by the purpose of providing an appropriate parametrization of the anisotropy profile, which is traditionally quantified by the so-called anisotropy parameter

$$\beta(r) = 1 - \frac{\sigma_\theta^2(r)}{\sigma_r^2(r)}, \tag{1}$$

where σ_θ and σ_r are dispersions of the tangential and radial velocity respectively. This part of the distribution function takes the following form

$$f_L(L) = \left(1 + \frac{L^2}{2L_0^2}\right)^{-\beta_\infty + \beta_0} L^{-2\beta_0}, \tag{2}$$

where β_0 and β_∞ are the asymptotic values of the anisotropy parameter at the halo centre and at infinity respectively. The scale of transition between these two asymptotes is determined by L_0, whereas a typical radial range of the growth or decrease of $\beta(r)$ is fixed at about 2 orders of magnitude centred on the radius corresponding to L_0 (see the anisotropy profiles in the top right panel of Fig. 1). Although some recent models of the distribution function offer a little more flexible parametrization of the anisotropy profile (e.g. Baes & Van Hese 2007; Van Hese et al. 2009), we find that our choice is quite suitable for the purpose of this work, given that we wish to reproduce the variability of $\beta(r)$ with as few parameters as possible.

The energy part of the distribution function $f_E(E)$ is given by the solution of the integral equation

$$\rho(r) = \iiint f_E(E) \left(1 + \frac{L^2}{2L_0^2}\right)^{-\beta_\infty + \beta_0} L^{-2\beta_0} d^3v. \tag{3}$$

This equation can be simplified to the one-dimensional integral equation and then solved numerically for f_E (see Appendix B in Wojtak et al. 2008). As an approximation for the density profile in (3) we use the NFW profile (Navarro, Frenk & White 1997), i.e.

$$\rho(r/r_s) = \frac{1}{4\pi(\ln 2 - 1/2)} \frac{1}{(r/r_s)(1 + r/r_s)^2} \frac{M_s}{r_s^3}, \tag{4}$$

where r_s is the scale radius and M_s is the mass enclosed in a sphere of this radius. Both parameters of the mass profile provide natural scales of the phase space coordinates so that any change of M_s or r_s corresponds to the expansion or contraction of a system in the space of velocities or the positions. For the sake of convenience, we use the scaling which fixes the range of positively defined binding energy per unit mass E at $[0, 1]$, namely r_s as the unit of radius and $V_s = (GM_s/r_s)^{1/2}(\ln 2 - 1/2)^{-1/2}$ as the unit of velocity (see Wojtak et al. 2008).

Due to projection effects, a fraction of the phase space is not accessible to observation. An observer is able to measure velocity along the line of sight $v_{\rm los}$ and the position on the sky which can be easily translated into the projected clustercentric distance R. This data set, when plotted $v_{\rm los}$ versus R, is commonly referred to as the phase space diagram or the velocity diagram. Data points in such diagrams are distributed according to the projected phase space density which is given by (e.g. Dejonghe & Merritt 1992; Merritt & Saha 1993; Mahdavi & Geller 2004)

$$f_{\rm los}(R, v_{\rm los}) = 2\pi R \int_{-z_{\max}}^{z_{\max}} {\rm d}z \iint_{E>0} {\rm d}v_R {\rm d}v_\phi f_E(E) f_L(L), \quad (5)$$

where z is the distance along the line sight from the cluster centre, while v_R and v_ϕ are velocity components in cylindrical coordinates. We will hereafter refer to $f_{\rm los}(R, v_{\rm los})$ as the phase space density. The integration area of the last two integrals is given by the circle $v_R^2 + v_\phi^2 < 2\Psi(\sqrt{R^2 + z^2}) - v_{\rm los}^2$, where $\Psi(r)$ is a positively defined gravitational potential for the NFW density profile, $\Psi(r) = V_s^2 \ln(1 + r/r_s)/(r/r_s)$ (Cole & Lacey 1996). The boundaries of the integral along the line of sight are given by the distance $\pm z_{\max}$ at which $v_{\rm los}(R)$ becomes the escape velocity. It is worth mentioning that in (5) we assume that the density profiles of a tracer (galaxies) and dark matter are proportional to each other. Second, $f_{\rm los}$ is defined up to the normalization which must therefore be imposed by the following additional condition

$$2 \int_0^{R_{\max}} {\rm d}R \int_0^{\sqrt{2\Psi(R)}} f_{\rm los}(R, v_{\rm los}) {\rm d}v_{\rm los} = 1, \qquad (6)$$

where $R_{\max}$ is a cut-off radius of the phase space diagram. From (6) one can immediately infer that the normalization factor is $\{r_s V_s [M_{\rm los}(R_{\max})/M_s]\}^{-1}$, where $M_{\rm los}(R_{\max})$ is the projected mass within the aperture $R_{\max}$ for the NFW density profile (see Łokas & Mamon 2001 for an analytical expression). From now on when referring to the phase space density $f_{\rm los}$ we will always mean that it is properly normalized as in (6).

We calculate the phase space density (5) numerically using the algorithm of Gaussian quadrature to evaluate each integral (Press et al. 1996). The energy part of the distribution function $f_E(E)$ is interpolated between points which are equally spaced in energy and provide the numerical solution of equation (3). In order to remove improper boundaries of the integral along the line of sight for $v_{\rm los} = 0$ we changed the variable z into Ψ. Next, to avoid singularity along the line $L^2 = (R^2 + z^2)v_\phi^2 + (v_R z - v_{\rm los} R)^2 = 0$, when $\beta_0 > 0$, we used an even number of abscissas for the variable v_ϕ so that $f_L(L)$ is never evaluated at $L = 0$.

Fig. 1 shows the contour maps of phase space density $f_{\rm los}(R, v_{\rm los})$ for 5 different anisotropy profiles plotted in the top right panel. In all cases we used $R_{\max} = 5r_s$, a typical virial radius of massive galaxy clusters, and the scale of transition between β_0 and β_∞, $L_0 = 0.2 V_s r_s$, which corresponds to $\sim 1 r_s$. To facilitate comparison, in the four bottom panels we plot the differences between the given and the isotropic ($\beta_0 = \beta_\infty = 0$) $f_{\rm los}$. The contours below the 0.01 threshold are neglected as insignificant for the typical number of data points considered in this work.

The results shown in Fig. 1 are quite intuitive to interpret. We notice that β_0 modifies $f_{\rm los}$ in the central part of the diagram in such a way that radially biased models predict more high velocity particles than tangentially biased ones (see the middle panels). On the other hand, β_∞ influences the outer part of a diagram so that the tangentially biased anisotropy suppresses $f_{\rm los}$ at $v_{\rm los} = 0$ and increases it for moderate velocities (see the bottom panels). In the limit of $\beta_\infty \ll 0$ the velocity distribution at $R/r_s \gg 1$ takes the characteristic horn-like shape with a minimum at $v_{\rm los} = 0$ and a maximum at the circular velocity $\pm[GM(R)/R]^{1/2}$ (see van der Marel & Franx 1993 for a qualitative comparison). As a final remark on Fig. 1, let us note that the parameters of the mass profile, M_s and r_s, manifest themselves only in the scaling of the axes of the diagram while preserving the shape of the phase space density. This property, when put together with the way how $f_{\rm los}$ depends on the anisotropy, automatically excludes the mass-anisotropy degeneracy that is an intrinsic problem of the analysis based on the Jeans equation (see e.g. Merrifield & Kent 1990).
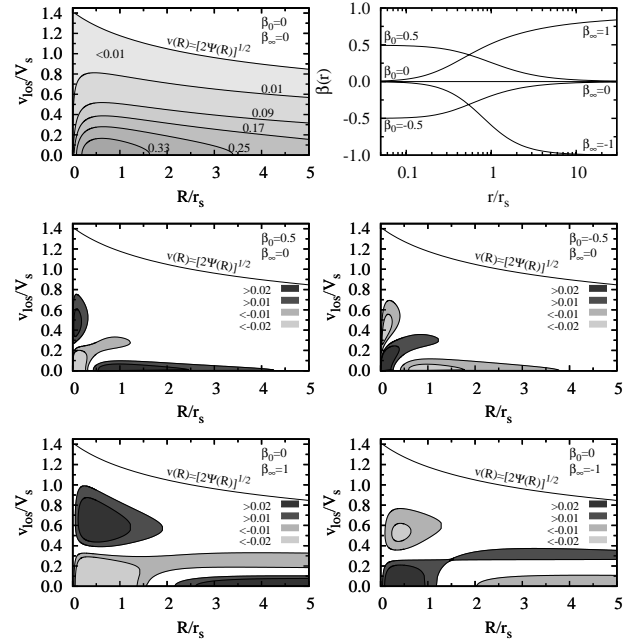

Figure 1. The projected phase space density $f_{\rm los}(R, v_{\rm los})$ for five anisotropy profiles characterized by different combinations of β_0 and β_∞ and plotted in the top right panel. The top left panel is a contour map of the isotropic $f_{\rm los}$ ($\beta_0 = \beta_\infty = 0$) and the four bottom panels show the differences between a given $f_{\rm los}$ and the isotropic one. The solid lines labelled by $v(R) = \sqrt{2\Psi(R)}$ are the profiles of the maximum escape velocity.

3 THE MOCK CATALOGUE OF PHASE SPACE DIAGRAMS

The model of the phase space density introduced in the previous section provides an idealized description of the real data. It neglects various secondary effects that occur in real galaxy clusters and perturb their phase space diagrams. Among those the most important seem to be: the breaking of spherical symmetry, the presence of substructures, the finite size of equilibrated zone and filamentary structures surrounding the clusters and infalling towards them. In order to study the impact of such effects on the reliability of
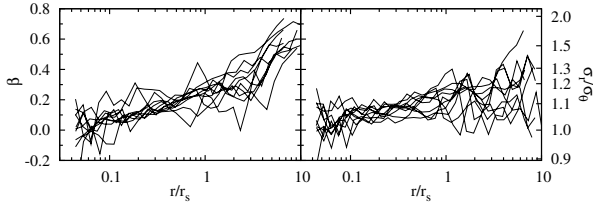
Figure 2. The anisotropy profiles in 20 dark matter haloes selected from the simulation. The two panels show separately the profiles for a subsample of ten haloes with more steeply (left) and less steeply (right) increasing profiles.

our approach to data analysis we need to construct a mock catalogue of phase space diagrams which on the one hand resembles the real spectroscopic cluster survey and on the other provides the true mass and anisotropy profiles. This can be easily achieved with the use of cosmological simulations.

We used an N-body cosmological simulation of a ΛCDM model. The simulation was carried out in a box of $160h^{-1}$Mpc with WMAP3 cosmological parameters (Spergel et al. 2007): $\Omega_m = 0.24$, $\Omega_\Lambda = 0.76$ and the dimensionless Hubble constant $h = 0.73$ (for more details see Wojtak et al. 2008). We identified all cluster-size haloes at redshift $z = 0$ and selected 20 of them that appeared not to be products of recent major mergers. For each halo we measured the virial mass M_v and the virial radius r_v which are defined in terms of the mean density within the virial sphere by the following equation

$$\frac{M_v}{(4/3)\pi r_v^3} = \Delta_c \rho_c, \tag{7}$$

where Δ_c is the so-called overdensity parameter. For cosmological model under consideration we adopted $\Delta_c = 94$ (see e.g. Łokas & Hoffman 2001). The virial mass of selected haloes is in the range $(1.5 - 20) \times 10^{14} M_\odot$. The minimum number of particles inside the virial sphere is 4.2×10^5.

By fitting the NFW formula (4) to the density profile calculated in radial bins equally spaced in logarithmic scale we determined the scale radius r_s of each halo and then the concentration parameter $c = r_v/r_s$. The scale mass M_s was measured directly as the mass enclosed by the sphere of radius r_s. We found that the mass profile extrapolated from r_s outward to the virial sphere recovers the virial mass with the precision of a few percent.

Fig. 2 shows the anisotropy profiles in selected dark matter haloes. The profiles were measured from velocity dispersions of dark matter particles inside thin spherical shells of radii extending to $1.5r_v$. Although a general trend for radial variation of the anisotropy is very clear, the profiles are considerably scattered. One can distinguish at least two classes of anisotropy profiles: steeply rising and flat or moderately rising. According to this criterion we divided our halo sample into two equally numerous groups of ten haloes each (see the left and right panel of Fig. 2). The reason for this division will become clear in section 5, where we consider constraints on the anisotropy profile from the joint analysis of many phase space diagrams.

In order to construct phase space diagrams associated with the haloes we pick a line of sight and find all particles

inside a cylinder of observation. The selection is restricted to the particles with the projected halocentric distances $R < r_v$ and velocities along the line of sight v_{los} (with the Hubble flow included) in the range ± 4000 km s^{-1} in the rest frame of a halo, which is the commonly used and sufficiently conservative velocity cut-off. The final phase space diagrams are generated by drawing randomly 300 particles for each halo, where we assume, to make the scheme unambiguous, that the tracer is given simply by the particles. The number of data points is fixed at a typical number of spectroscopic redshifts available for a nearby galaxy cluster with the same criterion of velocity cut-off. The selection of particles is done three times for three lines of sight chosen to be parallel to the Cartesian axes of the simulation box. Thus we effectively get three independent sets of phase space diagrams for the same sample of 20 haloes.

Due to projection effects, the phase space diagrams include the so-called interlopers: the particles (galaxies) of background or foreground which are not physically connected to the haloes (clusters). It is commonly accepted that any dynamical analysis should be necessarily preceded by an identification and removal of as many of these objects as possible. In this work we apply the procedure proposed by den Hartog & Katgert (1996) which appears to be one of the most effective algorithms for interloper removal (Wojtak et al. 2007; Wojtak & Łokas 2007). The method is based on an iterative rejection of galaxies with velocity exceeding a maximum velocity which is a function of the clustercentric distance R. The maximum velocity is calculated using a model which allows the galaxies to be either on circular orbit with velocity $v_{cir} = \sqrt{GM(r)/r}$ or to fall towards the cluster centre with velocity $\sqrt{2}v_{cir}$, where $M(r)$ is approximated in each iteration by the virial mass estimator (Heisler, Tremaine & Bahcall 1985). The phase space diagrams preprocessed in such a way are suitable for the proper analysis in terms of the phase space density. Hereafter, when referring to the phase space diagram we will always mean the diagram after interloper removal.

4   BAYESIAN ANALYSIS AND MCMC

Our aim is to constrain the parameters of the mass and anisotropy profiles by analyzing the mock phase space diagrams generated from the simulation. Following the principle of Bayesian inference our knowledge of model parameters $\boldsymbol{a}$ given the data D may be expressed as the posterior probability $p(\boldsymbol{a}|D)$ which is related to the probability of obtaining the data in a given model $p(D|\boldsymbol{a})$ (likelihood) and the prior probability $p(\boldsymbol{a})$ via equation

$$p(\boldsymbol{a}|D) = \frac{p(\boldsymbol{a})p(D|\boldsymbol{a})}{p(D)}, \tag{8}$$

where $\boldsymbol{a}$ is a vector of model parameters and D stands for the data. The probability $p(D)$ plays the role of the normalization coefficient and can be neglected without loss of generality.

The data D of a phase space diagram consists of N points located at $(R_i, v_{los,i})$, where $i = 1, \ldots, N$. Assuming that the number and mass density are proportional to each other the likelihood is given by

$$p(D|\boldsymbol{a}) = \prod_{i=1}^{N} f_{\mathrm{los}}(R_i, v_{\mathrm{los},i}|\boldsymbol{a}), \tag{9}$$

where f_{los} is the phase space density given by (5) and $\boldsymbol{a} = \{M_s, r_s, \beta_0, \beta_\infty, L_0\}$. It is worth mentioning that this formula does not take into account statistical errors of v_{los} and R. This is motivated by the fact that the errors of spectroscopic redshift measurements for nearby clusters are quite small and have negligible impact on the results of the analysis (see e.g. van der Marel et al. 2000).

4.1 Priors and reparametrization

The two parameters of the mass profile, M_s and r_s, are the scale parameters. The most appropriate prior probability for such parameters is the Jeffreys prior which represents equal probability per decade (Jeffreys 1946), i.e. $p(M_s) \propto 1/M_s$ and $p(r_s) \propto 1/r_s$. We will use this kind of prior in the analysis of any single phase space diagram. The situation will change in subsection 5.3 where we consider joint analysis of many phase space diagrams.

The commonly used anisotropy parameter $\beta(r)$ gives rise to unequal weights for tangentially and radially biased models. In order to make the prior probability for both types of anisotropy equal we introduce, following Wilkinson et al. (2002) and Mahdavi & Geller (2004), the following reparametrization

$$\ln \frac{\sigma_r(r)}{\sigma_\theta(r)} = -\frac{1}{2} \ln[1 - \beta(r)]. \tag{10}$$

Using a uniform prior for $\ln(\sigma_r/\sigma_\theta)_0 = -1/2 \ln(1 - \beta_0)$ and $\ln(\sigma_r/\sigma_\theta)_\infty = -1/2 \ln(1 - \beta_\infty)$ we put equal weights to all types of anisotropy. The ranges of both priors are limited by $\beta_0 \leq 1/2$ and $\beta_\infty \leq 0.99$. The first condition is motivated by the requirement of the distribution function to be positive (see An & Evans 2006), whereas the second one allows us to avoid improper posterior distribution that is important when the data favour $\beta_\infty \lesssim 1$.

The scale of transition between the two asymptotic values of anisotropy is defined by L_0. We find that keeping this parameter free causes degeneracy with β_∞ and β_0. This considerably restricts the information on the growth or decrease of the anisotropy profile that β_0 and β_∞ are expected to convey. To overcome this difficulty we fix L_0 at $0.2V_s r_s$, a value corresponding to the $\sim 1r_s$ transition scale expected from cosmological simulations (see Wojtak et al. 2008). This implies that the anisotropy profiles approach their asymptotic values at some characteristic scales of radius, namely $\beta_0 = \beta(r \lesssim 0.1r_s)$ and $\beta_\infty = \beta(r \gtrsim 10r_s)$. Fixing L_0 reduces the dimension of the parameter space by one so that the analysis is based on the four-parameter phase space model specified by M_s, r_s, β_0 and β_∞.

4.2 MCMC approach

Our purpose is to calculate the posterior probability density and provide credibility regions in the parameter space. This involves numerical integrations of multivariate semi-Gaussian functions that can be efficiently tackled with the Markov Chain Monte Carlo (MCMC) algorithm. The main idea of this approach is to construct a sufficiently long chain of models which are distributed in the parameter space according to the posterior probability density. Once such a chain is provided, one can easily compute the marginal probability distributions by projecting all points on to an appropriate parameter subspace and evaluating the histograms.

We construct the Markov chains following the Metropolis-Hastings algorithm (see e.g. Gregory 2005). In the first step of this algorithm one picks a trial point in parameter space $\boldsymbol{a}_{t+1}$ by drawing from the so-called proposal distribution $q(\boldsymbol{a}|\boldsymbol{a}_t)$ centred on the previous point of the chain $\boldsymbol{a}_t$. Then the point $\boldsymbol{a}_{t+1}$ is accepted with the probability equal to $\min\{1, r\}$, where r is the so-called Metropolis ratio given by

$$r = \frac{p(\boldsymbol{a}_{t+1}|D)q(\boldsymbol{a}_t|\boldsymbol{a}_{t+1})}{p(\boldsymbol{a}_t|D)q(\boldsymbol{a}_{t+1}|\boldsymbol{a}_t)}, \tag{11}$$

otherwise it takes the value of its predecessor, $\boldsymbol{a}_{t+1} = \boldsymbol{a}_t$. In the case of a few parameter model the proposal probability density can be any function that roughly resembles the target posterior distribution. In our work we use a multivariate Gaussian with a diagonal covariance matrix. The variances are calculated from short trial Markov chains of about 2000 models, where our best initial guess for parameter variances is used. This approach is a simplified version of the scheme outlined in Widrow, Pym & Dubinski (2008). It is worth mentioning that the proposal distribution is symmetric, $q(\boldsymbol{a}_{t+1}|\boldsymbol{a}_t) = q(\boldsymbol{a}_t|\boldsymbol{a}_{t+1})$, so that the Metropolis ratio is given by $p(\boldsymbol{a}_{t+1}|D)/p(\boldsymbol{a}_t|D)$.

An important issue of the MCMC analysis is a relative number of distinct points in the parameter space which is described by the so-called acceptance rate (the average rate at which proposed models are accepted). A recommended value of this parameter for many-parameter models is around $20 - 30$ percent (see e.g. Gregory 2005). We note that our choice of the proposal distribution keeps the acceptance rates of the resulting Markov chains within this range. Some additional modification of the proposal distribution is required in subsection 5.3, where we consider a joint analysis involving $N_p = 22$ parameters. In order to keep the acceptance rate at a desirable level all initial variances of the proposal distribution are scaled by $2.4^2/N_p$ (see e.g. Gelman et al. 1995). This correction prevents the acceptance rate from dropping, due to a large number of parameters in use, to a very low value of around 1 percent.

When applying the MCMC algorithm, it is crucial to check whether the chains explore properly the parameter space, i.e. whether they possess the property often referred to as mixing. To save computing time, we decided not to create several chains for each data sample that is a commonly advised way to look for convergence. As an alternative, we follow two simple indicators of mixing. First, we calculate parameter dispersions in two halves of a given chain and within the whole chain. If the relative differences between them do not exceed 10 percent for each parameter, the chain is considered to be mixed. Second, we monitor the variation of the posterior probability along the chains. Unless the profiles of $p(\boldsymbol{a}|D)$ exhibit a long-scale tendency to grow or decline, the chain is again expected to be mixed. We note that the so-called burn-in part of the chains, when the first models gradually approach the most favoured zone of the parameter space, is not longer than 1 percent of the total length of the chains. All chains used in this analysis consist
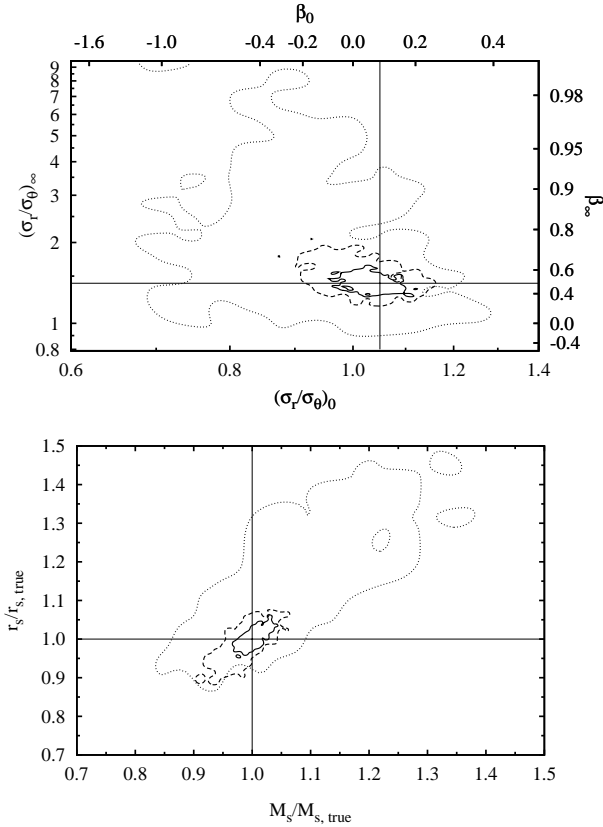
Figure 4. Projected phase space diagram obtained from the full phase space distribution function with anisotropy parameters $\beta_0 = 0.1$ and $\beta_\infty = 0.5$. Solid and dashed lines show respectively the isodensity contours of the best-fitting model and the smoothed surface density of 9000 sampling points. Two envelope lines labelled by $v(R) = \pm\sqrt{2\Psi(R)}$ are the profiles of the escape velocity.

Figure 3. The 1σ credibility regions inferred from the MCMC analysis of three theoretical phase space diagrams generated from the distribution function with $\beta_0 = 0.1$ and $\beta_\infty = 0.5$. The parameters of the mass profile were rescaled by their true values. Solid, dashed and dotted lines correspond to the diagrams with 9000, 3000 and 300 data points. Straight lines indicate the true values of the parameters.

of 10^4 models that is above the recommended minimum of $330N_p$ (Dunkley et al. 2005).

4.3 Number of data points

One of the leading factors affecting the posterior probability is the limited number of data points. In order to study the impact of this effect we carried out the analysis of three theoretical phase space diagrams with 300, 3000 and 9000 points that correspond to a typical size of a data sample for a single nearby galaxy cluster (the first number) and to a compilation of $10 - 30$ of them (the last two numbers). The diagrams were generated from a discrete representation of the full phase space distribution function with the following parameters of the anisotropy profile: $\beta_0 = 0.1$, $\beta_\infty = 0.5$ and $L_0 = 0.2V_s r_s$. The phase space was sampled using the acceptance-rejection technique (Press et al. 1996) in a manner described by Kazantzidis, Magorrian & Moore (2004).

Fig. 3 shows the 1σ credibility regions, i.e. the regions enclosing 68.3 percent of the corresponding marginal probability, inferred from the MCMC analysis of the three theoretical diagrams. For the sake of simplicity, the parameters of the mass profile were rescaled by the true values. The best-fitting model for $N = 9000$ case is overplotted in solid lines
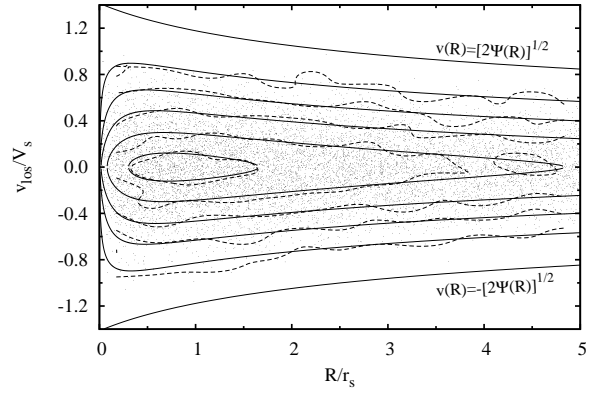
on top of the smoothed contour map of the phase space density (dashed lines) in Fig. 4. From Fig. 3 we conclude that typical relative errors of M_s and r_s from the analysis of a single phase space diagram with $N = 300$ points are about 20 percent. On the other hand, the corresponding constraints on the anisotropy parameters are very poor. The marginal distribution is so wide, particularly along β_∞, that it is expected to be sensitive to any kind of noise in the real data. Anticipating the results of the following section we emphasize that the only way to reliably constrain the anisotropy profile is to increase the number of data points. In practice, this can be achieved in the joint analysis of at least 10 independent diagrams (3000 data points), where one assumes a universal anisotropy profile. The result for $N = 3000, 9000$ in the top panel of Fig. 3 shows what we can expect from such analysis.

5 RESULTS OF THE ANALYSIS

We carried out the MCMC analysis of 60 phase space diagrams of our mock data catalogue described in section 3. The results are presented in the form of the maximum a posteriori (MAP) values of the parameters and the errors that correspond to the boundary of the 1σ credibility regions of the marginal distributions, i.e. the regions enclosing 68.3 percent of the marginal probability.

5.1 Mass profile

Fig. 5 shows the results for the parameters of the mass profile, M_s and r_s. From the diagrams showing residuals, attached below the mains plots, we conclude that in half of the cases the relative errors do not exceed 25 percent (see the lines of the first and third quartile of the residuals). On the other hand, around 15 percent of the results differ from the real values by more than 50 percent that can hardly be reconciled with the expectations from the effect of a shot
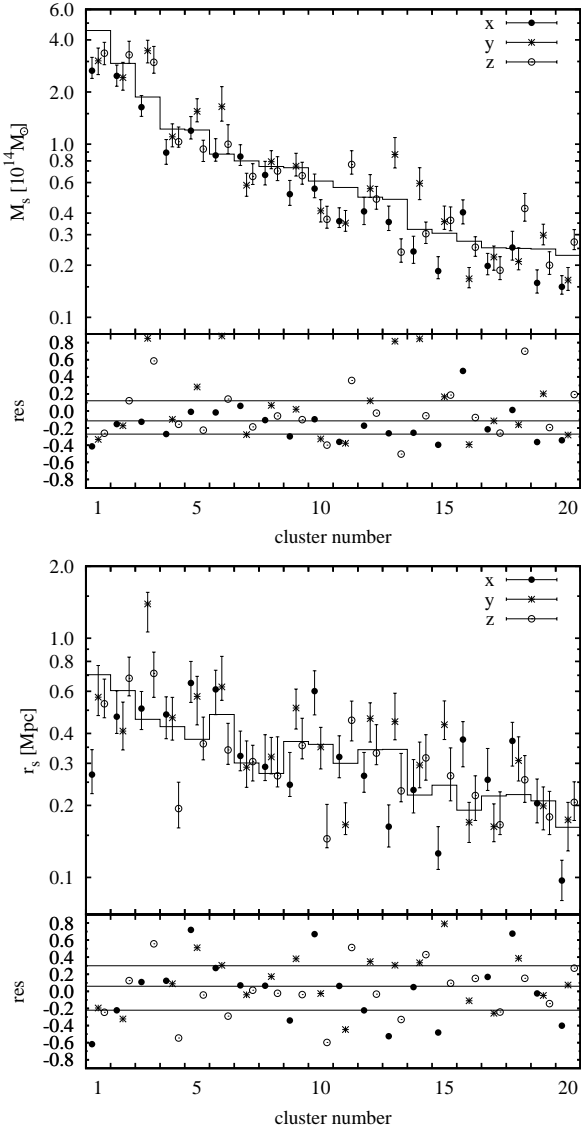
Figure 5. Mass profile parameters obtained from the MCMC fits of 20 individual simulated clusters observed in 3 directions. The symbols indicate MAP (maximum a posteriori) values and the errors correspond to the boundary of the 1σ credibility region of the marginal distributions. Filled circles, asterisks and empty circles refer to directions of projection along the x, y and z axis of the simulation box. Solid broken lines indicate the true values of the parameters. The three lines in the lower panels showing fractional residuals from the true parameter values indicate the quartiles of a total set of 60 residuals. The clusters are ordered by decreasing scale mass M_s.

Figure 6. Same as Fig. 5, with the scale mass and radius converted to the virial mass M_v and concentration c. The clusters are ordered by decreasing virial mass.

rameter estimates and the halo ellipticity or the alignment of the major axis with respect to the line of sight.

All Markov chains can be easily transformed into the chains of models parametrized by the virial mass M_v and the concentration c. The resulting credibility ranges and the MAP values of these parameters are shown in Fig. 6. We find that the scatter in residuals is comparable to the case where the scale parameters were used and the positions of the most discrepant points imply the presence of similar systematic errors. There is noticeable progress in the constraints on the concentration parameter in comparison to previous work based on modelling velocity moments (Sanchis et al. 2004; Łokas et al. 2006). On the other hand, there is no convincing evidence for a similar improvement for the virial mass.

The MAP values of the mass parameters typically underestimate the true values by 15 percent and 11 percent for the virial and scale mass respectively (see the position of me-

noise. Analyzing a few tens of theoretical phase space diagrams from subsection 4.3 with $N = 300$ data points we find that residuals induced by a shot noise do not exceed 30 percent. This may suggest that the most outlying points in Fig. 5 are probably subject to non-negligible systematic errors. Potential sources of these errors could include projection effects as well as those associated with internal structure. We looked for the influence of the shape of dark matter haloes and found no correlation between the accuracy of pa-
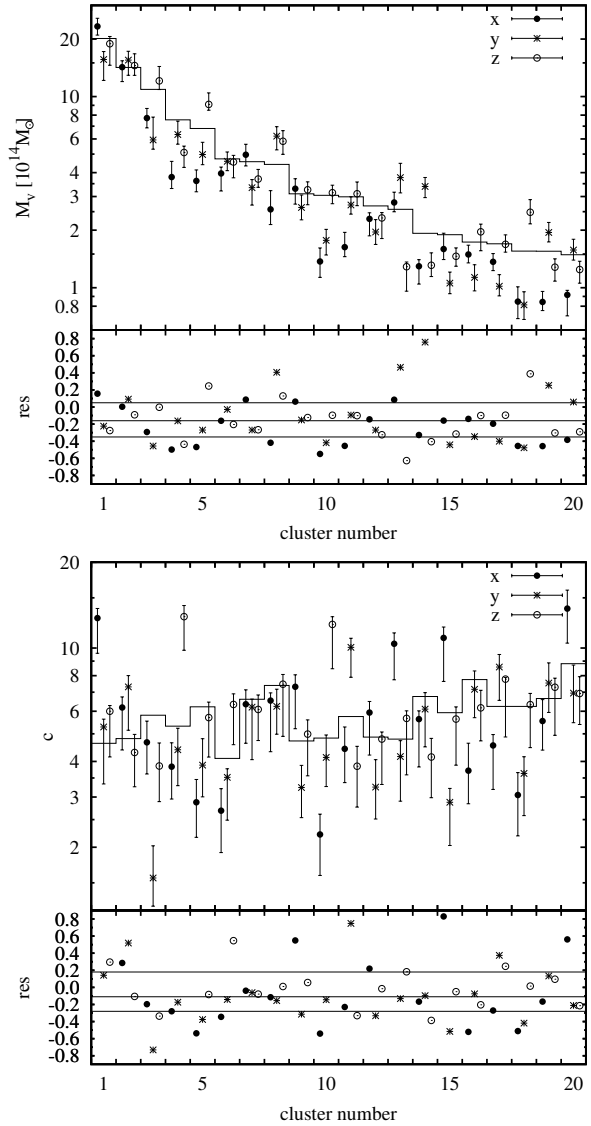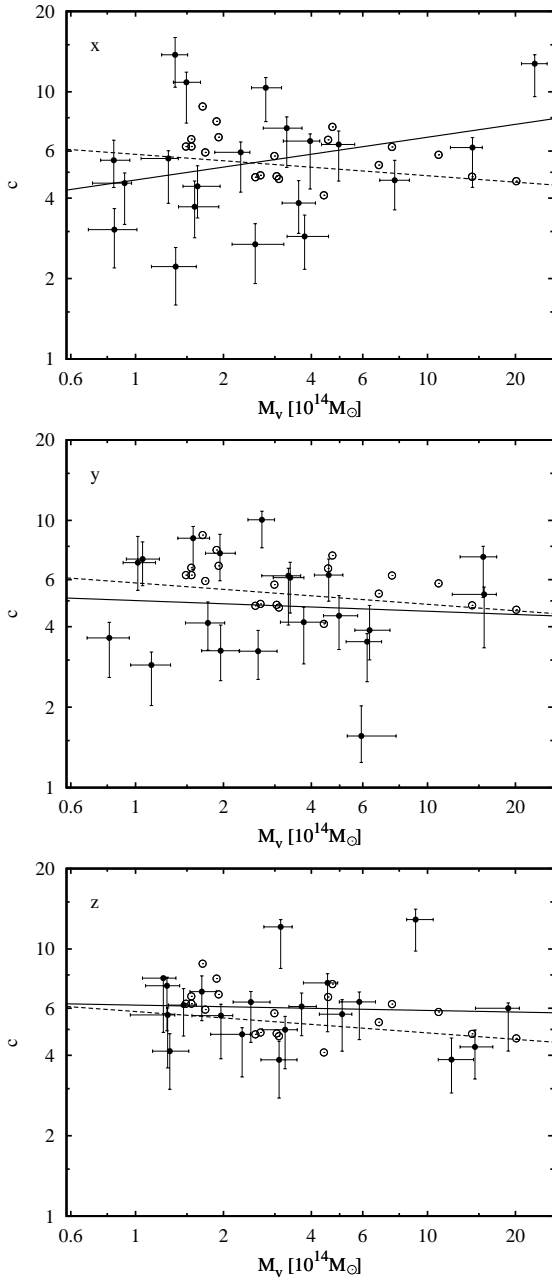
Figure 7. Prospects of reconstructing the mass-concentration relation, based on our MCMC analysis of 20 individual simulated clusters observed in 3 directions labelled by x, y and z (from the top to the bottom panel). The solid lines are the power-law fits to the best-fitting parameters obtained from the analysis of the phase space diagrams (filled circles). Open circles indicate the true values of the parameters for our 20 clusters. The dashed line repeated in each panel is a power-law fit to the mean mass-concentration relation for WMAP3 cosmology from Macciò et al. (2008).

dian lines in the diagrams showing residuals in Figs. 6 and 5). Interestingly, a similar offset was reported by Biviano et al. (2006) for the virial mass estimated from the Jeans analysis of velocity dispersion profiles. They suggested that this bias is related to the presence of interlopers infalling towards the cluster along filaments. Due to relatively small

velocities in the rest frame of a cluster, these objects cannot be identified by any algorithm of interloper removal and, therefore, remain in the sample decreasing the velocity dispersion and the resulting virial mass. We think that a similar mechanism is likely responsible for the offset of our results. Nevertheless, we emphasize that this bias is smaller than the statistical errors obtained in the MCMC analysis so that the overall effect is not statistically significant.

It is a well known fact the concentration parameter is weakly correlated to the virial mass (e.g. Navarro et al. 1997; Bullock et al. 2001). This so-called mass-concentration relation is an imprint of the formation history and is therefore thought to be one of the predictions of the cosmological model. Fig. 7 demonstrates the prospects of recovering this relation with our approach. Each panel in this Figure shows the results obtained for 20 phase space diagrams for a given direction of observation (filled circles with error bars) and the best power-law fit (solid line). For comparison with the prediction of the ΛCDM model, we plot the true parameters of our 20 clusters (open circles) and with a dashed line the power-law fit to the mean mass-concentration relation for relaxed haloes simulated in the framework of WMAP3 cosmology from Macciò, Dutton & van den Bosch (2008). We find that, within errors, the empirical M_v-c relation is consistent with a flat profile in all cases. For one projection (top panel) the slope deviates from the expected value of -0.1 by more than two sigma. The normalization within the mass range under consideration does not exhibit any offset with respect to the prediction so that future determination of this quantity for real clusters seems to be feasible. This result is particularly important in the context of recently discussed inconsistency between the normalization from the simulations and observational constraints (see e.g. Comerford & Natarajan 2007).

5.2 The anisotropy profile

Fig. 8 shows the results of the MCMC analysis for the anisotropy parameters. Since the phase space model is well-defined only within the virial sphere, we decided to replace the parameter β_∞ with β_{5r_s} measured at $5r_s$ which is the typical scale of the virial radius. This allows us to avoid extrapolation of the model beyond the virial sphere. For comparison the Figure also shows in gray boxes the range of anisotropy values at $0.1r_s$ and $5r_s$ in two samples of clusters introduced is section 3 which differ in steepness of their anisotropy profiles.

The majority of results for β_0 are consistent with the true values. The errors are quite large, however, so they would not allow to distinguish between radially and tangentially biased anisotropy. We emphasize that all 1σ credibility ranges are clearly detached from the upper boundary of the prior probability, $\beta = 0.5$. This means that this limit is not artificial but is embedded in the phase space structure of an equilibrated system.

The anisotropy at $5r_s$ is poorly constrained. Although there is a weak tendency for the results to cluster around the true values, most of them are shifted considerably upwards or downwards. An important circumstance responsible for this bias is the fact that the posterior probability distribution is so wide in the space of anisotropy parameters that it is sensitive to any irregularities in the phase space dia-
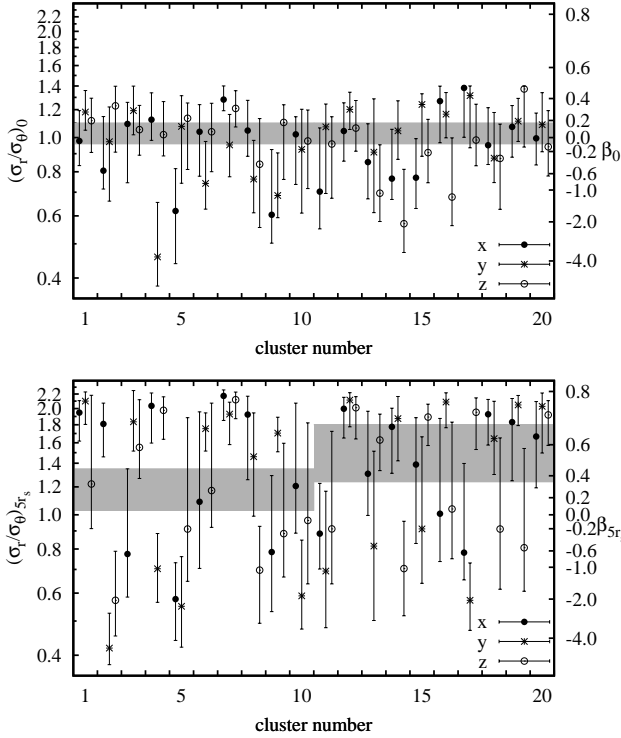
Figure 8. Anisotropy profile parameters obtained from the MCMC fits of 20 individual simulated clusters observed in 3 directions. Gray boxes indicate the range of the true anisotropy at $0.1r_s$ (top panel) and $5r_s$ (bottom panel). The first and second ten clusters correspond to the samples with flat to moderately rising and steeply rising anisotropy profiles respectively, as shown in Fig. 2. The meaning of all symbols is the same as in Fig. 5.

gram. These irregularities often occur in the outermost part of the diagrams, mostly due to subclustering and the presence of small velocity interlopers from the outside of the virial sphere. The overall consequence is that the posterior probability is often perturbed in β_∞ (or β_{5r_s}) and typically stable in β_0. Due to significant systematic errors of β_{5r_s} (or β_∞), the constraints on the radial variation of the anisotropy for a single galaxy cluster are very weak. Most probably this caveat cannot be avoided in any approach aiming to determine the anisotropy profile of galaxies in a cluster (see e.g. Hwang & Lee 2008).

5.3 The anisotropy profile from the joint analysis

The only possibility to obtain reliable constraints on the anisotropy profile is to increase the number of data points used in the analysis. In subsection 4.3 we showed that a sample of 3000 data points is expected to provide conclusive results in this matter. When dealing with galaxy clusters, such increase in the amount of data can only be achieved at present by means of the joint analysis of several clusters (~ 10 clusters). An additional advantage of this approach is that all local irregularities of the phase space diagrams compensate each other minimizing systematic errors.

A common approach to carry out a joint analysis of several clusters consists of two stages. First, one merges all phases space diagrams with properly rescaled radii R and
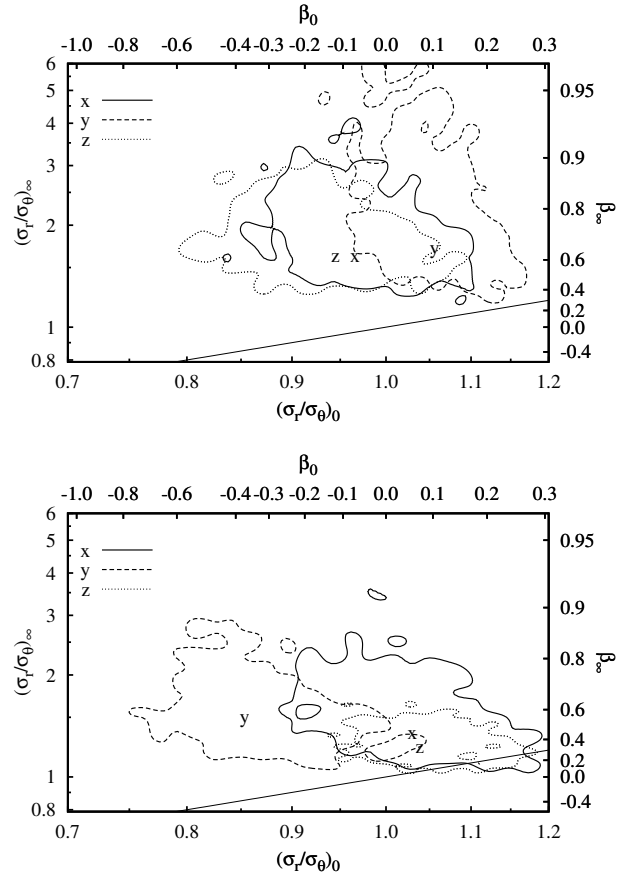

Figure 9. Anisotropy profile parameters obtained from the joint MCMC fits of 10 simulated clusters with rapidly (top panel) and moderately (bottom panel) rising anisotropy profiles. The contours indicate the boundaries of the 1σ credibility regions with solid, dashed and dotted lines referring to the projection along the x, y and z axis respectively. The corresponding maximum a posteriori (MAP) values are indicated with letters x, y and z. The straight solid line represents a family of flat anisotropy profiles.

velocities $v_{\rm los}$. As the units of phase space coordinates one most often uses the virial radius r_v and the so-called virial velocity $V_v = (GM_v/r_v)^{1/2}$. Then, assuming homology of the cluster sample, the resulting phase space diagram is reanalyzed in the framework of a model reduced by the parameters that were used to scale the radii and velocities. The final constraints obtained in this approach are thought to represent typical properties of galaxy clusters in a given sample (e.g. van der Marel et al. 2000; Mahdavi & Geller 2004; Łokas et al. 2006).

Our approach differs from the above scheme in both respects. First, instead of merging many phase space diagrams we introduce the following generalization of the likelihood function (9) for the joint analysis of n phase space diagrams

$$f(M_{s,1}, \ldots, M_{s,n}, r_{s,1}, \ldots, r_{s,n}, \beta_0, \beta_\infty) =$$
$$\prod_{j=1}^{n} \prod_{i=1}^{N} f_{\rm los}(R_{j,i}, v_{{\rm los},j,i} | \{M_{s,j}, r_{s,j}, \beta_0, \beta_\infty\}), \qquad (12)$$

where j and i are respectively the reference number of a cluster and a data point of j-th phase space diagram. The likelihood as well as the posterior probability distribution
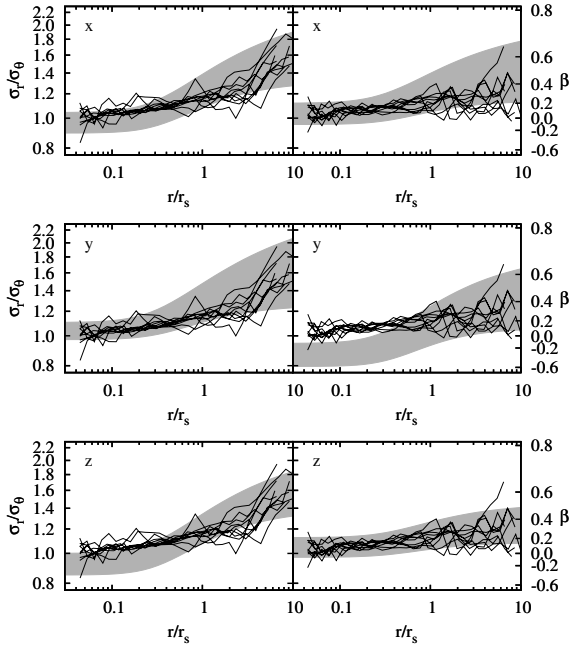
Figure 10. Anisotropy profile obtained from the joint MCMC fits of 10 simulated clusters with rapidly (left panels) and moderately (right panels) rising anisotropy profiles, in three directions of observation (rows). The widths of the shaded regions are given by the 1σ credibility range of the anisotropy at a given radius. Solid lines are the true anisotropy profiles measured from the simulations.

are functions of $2n + 2$ parameters (the parameter L_0 is fixed as discussed in subsection 4.1). We assume that the two parameters of the anisotropy profile are common to all clusters so that the final result concerning both of them is expected to provide a typical variation of the anisotropy in a given cluster sample. Our knowledge of $M_{s,i}$ and $r_{s,i}$ obtained in the previous analysis (see Fig. 5) is incorporated in the prior probability. For the sake of maintaining the analytical form of the prior, we used the product of $2n$ Gaussian distributions centred at the MAP values of $M_{s,i}$ and $r_{s,i}$. The dispersion of each Gaussian was fixed at the dispersion of the corresponding marginal probability distribution. It is interesting to note that if we arbitrarily ignored the credibility region of the nuisance parameters $M_{s,i}$ and $r_{s,i}$, our approach would be conceptually very similar to the case when all phase space diagrams are merged except that radii and velocities would be scaled by r_s and V_s.

We analyzed two samples of clusters, with $n = 10$ objects each, observed in three directions as described in section 3. The samples separate the objects with distinctly rising anisotropy profiles from those with flat or moderately rising ones. This provides an opportunity to verify whether our approach is able to distinguish between these two cases. Fig. 9 shows the 1σ credibility regions in the $\beta_0 - \beta_\infty$ plane inferred from the Markov chains of 10^4 models. First, we notice that all results are consistent with the velocity dispersion tensor that is isotropic in the cluster centre and radially biased outside. Second, the resulting anisotropy profiles are clearly steeper for clusters with rising anisotropy profiles (top panel) and flatter for the others (bottom panel).

In order to check the radial variation of the anisotropy

in Fig. 10 we plot the 1σ credibility regions of its radial profiles. The resulting anisotropy effectively traces the true profiles (solid lines) within the radial range covering more than two orders of magnitude around r_s. A local deviation from the true values occurs only once (middle right panel for $r < 0.5r_s$) as a result of relatively low quality of this data sample (see the results shown with asterisks in Fig. 8). Since the constraints on β_0 are considerably tighter than on β_∞, statistical errors of local anisotropy typically increase with radius. Nevertheless, it is still possible to distinguish between steeply rising (left panels) and flat (right panels) profiles. Note that the accuracy achieved in this analysis is close to the upper limit, since statistical errors become comparable to the internal dispersion of the anisotropy profiles in the cluster samples (see e.g. the bottom right panel of Fig. 10). Still, one may still expect to improve the results and minimize systematic errors by adding more clusters.

6   DISCUSSION

We have performed the Bayesian analysis of mock phase space diagrams of galaxy clusters in terms of a fully anisotropic model of the phase space density. Our approach allows to constrain parameters of the total mass profile, M_s and r_s, as well as the asymptotic values of the anisotropy profile, β_0 and β_∞. The phase space model was designed to detect the radial variation of the anisotropy profile within a fixed distance range covering two orders of magnitude in radius around r_s. The choice of this radial range is motivated by the results from cosmological simulations (Wojtak et al. 2008) and our goal to avoid degeneracy between radial scales and the asymptotes of the anisotropy.

Parameters of the mass profile are determined in our approach with rather satisfying average accuracy of about 30 percent. On the other hand, around 15 percent of the results are probably subject to systematic error and differ from the true values by $50 - 90$ percent. Nevertheless, it is worth noting that the typical relative errors of 30 percent in the virial mass are comparable to the potential accuracy of mass determination from other methods, such as modelling of X-ray gas (e.g. Nagai, Vikhlinin & Kravtsov 2007), analysis of velocity moments (Sanchis, Łokas & Mamon 2004) or the standard virial mass estimator (e.g. Biviano et al. 2006). The constraints on the concentration parameter are however more reliable and tighter in comparison to the approach based on velocity moments (Sanchis, Łokas & Mamon 2004; Łokas et al. 2006).

We find that the scale mass M_s and the virial mass M_v tend to be underestimated on average by 11 and 15 percent respectively. It is very likely that this bias is caused by small velocity interlopers, as described in Biviano et al. (2006), which are not tractable by algorithms for interloper removal and effectively decrease the velocity dispersion of a system. Interestingly, a similar offset in the virial mass estimates is noticed in the analysis of mock X-ray data of clusters (e.g. Ameglio et al. 2009; Nagai et al. 2007). However, this happens probably incidentally, since these authors claim that their bias is due to the lack of hydrostatic equilibrium in the outer parts of clusters.

The constraints on the anisotropy profile of a single cluster are barely conclusive. The reason for this is the pres-

ence of substructures at large clustercentric distances which gives rise to significant systematic errors of β_∞. The final effect is that, although the central asymptotic anisotropy is determined rather precisely, the overall constraint on the anisotropy profile is rather poor. This situation changes in the case of a joint analysis of several phase space diagrams. We find that it is then easy to obtain reliable constraints on the radial variation of the anisotropy within the radial range of two orders of magnitude around the scale radius r_s. Note that we adopted the phase space scaling that is consistent with the intrinsic parameters of the mass profile: M_s and r_s, in contrast to commonly used assumption that a sample of phase space diagrams is homologous with respect to scaling by the virial radius and virial velocity (e.g. van der Marel et al. 2000; Łokas et al. 2006).

In this work we demonstrated the potential and discussed the reliability of the analysis of kinematic data of galaxy clusters in terms of an anisotropic phase space density model. The method is able to provide robust constraints on the parameters of the total mass profile and the mean anisotropy profile. The results could certainly contribute to tests of the mass-concentration relation as well as improve constraints on the mean anisotropy profile of galaxy clusters (e.g. van der Marel et al. 2000; Biviano & Katgert 2004; Łokas et al. 2006). This will be the subject of a follow-up work where we analyze real kinematic data for a sample of ~ 20 nearby galaxy clusters.

REFERENCES

Ameglio S., Borgani S., Pierpaoli E., Dolag K., Ettori S., Morandi A., 2009, MNRAS, 394, 479
An J. H., Evans N. W., 2006, ApJ, 642, 752
Ascasibar Y., Gottlöber S., 2008, MNRAS, 386, 2022
Baes M., van Hese M., 2007, A&A, 471, 419
Benatov L., Rines K., Natarajan P., Kravtsov A., Nagai D., 2006, MNRAS, 370, 427
Binney J., Mamon G. A., 1982, MNRAS, 200, 361
Binney J., Tremaine S., 2008, Galactic Dynamics. Princeton Univ. Press, Princeton
Biviano A., Katgert P., 2004, A&A, 424, 779
Biviano A., Murante G., Borgani S., Diaferio A., Dolag K., Girardi M., 2006, A&A, 456, 23
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Chanamé J., Kleyna J., van der Marel R., 2008, ApJ, 682, 841
Cole S., Lacey C., 1996, MNRAS, 281, 716

Comerford J. M., Natarajan P., 2007, MNRAS, 379, 190
Dejonghe H., Merritt D., 1992, ApJ, 391, 531
den Hartog R., Katgert P., 1996, MNRAS, 279, 349
Dunkley J., Bucher M., Ferreira P. G., Moodley K., Skordis C., 2005, MNRAS, 356, 925
Gelman A., Carlin J. B., Stern H. S., Rubin D. B., 1995, Bayesian Data Analysis. Chapman & Hall, London
Gerhard O., Jeske G., Saglia R. P., Bender R., 1998, MNRAS, 295, 197
Gregory P. C., 2005, Bayesian Logical Data Analaysis for the Physical Science. Cambridge Univ. Press, Cambridge
Heisler J., Tremaine S., Bahcall J. N., 1985, ApJ, 298, 8
Hwang H. S., Lee M. G., 2008, ApJ, 676, 218
Jeffreys H., 1946, Proc. Roy. Soc. A, 186, 453
Kazantzidis S., Magorrian J., Moore B., 2004, ApJ, 601, 37
Kent S. M., Gunn J. E., 1982, AJ, 87, 945
Łokas E. L., 2002, MNRAS, 333, 697
Łokas E. L., Hoffman Y., 2001, in Spooner N. J. C., Kudryavtsev V., eds, Proc. 3rd International Workshop, The Identification of Dark Matter. World Scientific, Singapore, p. 121
Łokas E. L., Mamon G. A., 2001, MNRAS, 321, 155
Łokas E. L., Mamon G. A., 2003, MNRAS, 343, 401
Łokas E. L., Wojtak R., Gottlöber S., Mamon G. A., Prada F., 2006, MNRAS, 367, 1463
Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, MNRAS, 391, 1940
Mahdavi A., Geller M. J., 2004, ApJ, 607, 202
Mamon G. A., Łokas E. L., 2005, MNRAS, 363, 705
Merrifield M. R., Kent S. M., 1990, AJ, 99, 1548
Merritt D., 1987, ApJ, 313, 121
Merritt D., Saha P., 1993, ApJ, 409, 75
Nagai D., Vikhlinin A., Kravtsov A. V., 2007, ApJ, 655, 98
Natarajan P., Kneib J.-P., 1996, MNRAS, 283, 1031
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1996, Numerical Recipes in Fortran 77. Cambridge Univ. Press, Cambridge
Sanchis T., Łokas E. L., Mamon G. A., 2004, MNRAS, 347, 1198
Schwarzschild M., 1979, ApJ, 232, 236
Solanes J. M., Salvador-Solé E., 1990, A&A, 234, 93
Spergel D. N., Bean R., Doré O. et al., 2007, ApJS, 170, 377
van der Marel R., Franx M., 1993, ApJ, 407, 525
van der Marel R. P., Magorrian J., Carlberg R. G., Yee H. K. C., Ellingson E., 2000, AJ, 119, 2038
Van Hese E., Baes M., Dejonghe H., 2009, ApJ, 690, 1280
Widrow L. M., Pym B., Dubinski J., 2008, ApJ, 679, 1239
Wilkinson M. I., Kleyna J., Evans N. W., Gilmore G., 2002, MNRAS, 330, 778
Wojtak R., Łokas E. L., 2007, MNRAS, 377, 843
Wojtak R., Łokas E. L., Gottlöber S., Mamon G. A., 2005, MNRAS, 361, L1
Wojtak R., Łokas E. L., Mamon G. A., Gottlöber S., Prada F., Moles M., 2007, A&A, 466, 437
Wojtak R., Łokas E. L., Mamon G. A., Gottlöber S., Klypin A., Hoffman Y., 2008, MNRAS, 388, 815